09/784787

cofC

PATENT
Attorney Docket No. 55491

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Patent No. 7,027,397

Confirmation No. 3031

Issued: April 11, 2006

Name of Patentee: Turner et al.
Patent Title: METHOD AND APPARATUS FOR ACCUMULATING AND DISTRIBUTING TRAFFIC AND FLOW CONTROL INFORMATION IN A PACKET SWITCHING SYSTEM

CERTIFICATE OF MAILING

I hereby certify that this paper is being deposited with the United States Postal Service on the date shown with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, Washington, D.C. 20231, **on October 24, 2006.**

Kirk D. Williams, Esq. 10-24-2006

REQUEST FOR CERTIFICATE OF CORRECTION OF PATENT FOR PATENT OFFICE MISTAKE (37 C.F.R. § 1.322)

Attn: Certificate of Correction Branch
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

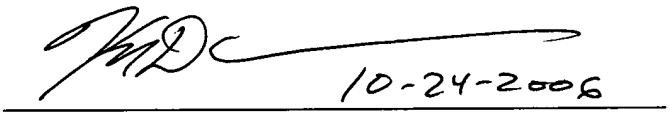
Dear Sir:

It is requested that a Certificate of Correction be issued to correct Office mistakes found the above-identified patent. Attached hereto is a Certificate of Correction which indicates the requested correction. For your convenience, also attached are copies of selected pages (a) from the issued patent with errors highlighted and (b) from the original application as filed February 15, 2001 with the correct text/instructions.

It is believed that there is no charge for this request because applicant or applicants were not responsible for such error, as will be apparent upon a comparison of the issued patent with the application as filed or amended. However, the Assistant Commissioner is hereby authorized to charge any fee that may be required to Deposit Account No. 501430.

Respectfully submitted,
The Law Office of Kirk D. Williams

Date: October 24, 2006

By  10-24-2006

Kirk D. Williams, Reg. No. 42,229
One of the Attorneys for Applicants
CUSTOMER NUMBER 26327
The Law Office of Kirk D. Williams
1234 S. OGDEN ST., Denver, CO 80210
303-282-0151 (telephone), 303-778-0748 (facsimile)

UNITED STATES PATENT AND TRADEMARK OFFICE

CERTIFICATE OF CORRECTION

PATENT NO. : 7,027,397

DATED : April 11, 2006

INVENTOR(S) : Turner et al.

It is certified that error(s) appear in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Col. 4, line 63, replace "213" with -- 2B --

Col. 7, line 46, replace "301 each" with -- 301, each --

Col. 7, line 48, replace "P/O" with -- I/O --

Col. 8, line 26, replace "P/O" with -- I/O --

MAILING ADDRESS OF SENDER:
Kirk D. Williams, Reg. No. 42,229
Customer No. 26327
The Law Office of Kirk D. Williams
P.O. Box 61538, Denver, CO 80206

PATENT NO. 7,027,397
No. of additional copies

⇨ NONE (0)

0 3 2006

described may include, inter alia, systems, integrated circuit chips, methods, and computer-readable medium containing instructions. The embodiments described hereinafter embody various aspects and configurations within the scope and spirit of the invention.

Accumulating and Distributing Flow Control Information

Methods and apparatus are disclosed for accumulating traffic information and distributing flow control information in a packet switching system. Traffic information is collected in multiple elements of the packet switching system. These multiple elements forward to collecting elements of the packet switching system indications of congestion and/or other types of information useful in determining traffic conditions within the packet switching system. The collecting elements manipulate the received indications of traffic conditions and generate flow control messages which are sent to individual sending components (e.g., I/O interfaces, line cards, ports, etc.) of the packet switching system.

In one embodiment, a switching element maintains for each destination a count of packets within itself which are addressed to the particular destination. Indications of this collected information are sent to all, or a subset of, the collecting switching elements of the packet switching system. These collecting elements accumulate the information received from multiple sources. The accumulated information is evaluated, and when a congestion condition is determined or anticipated, then flow control messages are distributed to all, or a subset of, the packet sources (e.g., I/O interfaces, line cards, ports, etc.).

In one embodiment, information is collected by a tabulator in each of the distribution switching elements of a switching fabric. For example, in one embodiment of a three stage switching fabric, a tabulator in each of the second stage switching elements maintains a tabulation data structure of packet counts corresponding to approximately to the number of packets currently within the particular switching element. Periodically or asynchronously, update information is sent from these tabulators to accumulators in one or more of the third stage switching elements. The accumulators then accumulate and manipulate the received traffic information and distribute flow control information to one or more components (e.g., I/O interfaces, line cards, ports, etc.) which can react to lessen the actual, perceived, or anticipated congestion or other traffic conditions. Traffic and flow control information may be sent using at least one of many different techniques such as in messages, over control lines, piggybacked in other data or control messages, etc.

Details of Exemplary Embodiments

FIGS. 1A–3C and their discussion herein are intended to provide a description of various exemplary packet switching systems. FIGS. 1A–C illustrate the basic topology of different exemplary packet switching systems. FIG. 1A illustrates an exemplary packet switch **100** having multiple inputs and outputs and a single interconnection network **110**. FIG. 1B illustrates an exemplary packet switch **140** having multiple interconnection networks **141** and folded input and output interfaces **149**. FIG. 1C illustrates an exemplary folded packet switch **160** having multiple interconnection networks **161** and folded input and output interfaces **169**. Embodiments of each of these packet switches **100**, **140** and **160** receive, generate, accumulate, distribute, and react to flow control information in the manners disclosed herein. Of course, the invention is not limited to these illustrated operating environments and embodiments, and the packet switching systems may have more or less elements.

FIG. 1A illustrates an exemplary embodiment of a packet switch **100**. Packet switch **100** comprises multiple input interfaces **105**, interconnection network **110**, and output interfaces **125**. Input interfaces **105** and output interfaces **125** are both coupled over multiple links to interconnection network **110**. Line cards **101** and **131** are coupled to input interfaces **105** and output interfaces **131**. In certain embodiments including other packet switching topologies, line cards or their functionality may be included in the packet switch itself, or as part of the packet switching system.

In one embodiment, interconnection network **110** comprises multiple switch elements SE-1 **112**, SE-2 **115**, and SE-3 **118** that are interconnected by multiple links. Line cards **101** and **131** may connect to other systems (not shown) to provide data items (e.g., packets) to be routed by packet switch **100**. Flow control information may be generated, consumed, or processed at one or more of the line cards **101**, **131**, input interfaces **105**, switch elements SE-1 **112**, SE-2 **115**, and SE-3 **118**, output interfaces **125**, and/or other locations within packet switch **100** or the packet switching system.

FIG. 1B illustrates another exemplary operating environment and embodiment of a packet switch **140**. Packet switch **140** comprises multiple folded input and output interfaces **149** interconnected over multiple links to interconnection networks **141**, which are interconnected over multiple links returning to input and output interfaces **149**. In one embodiment, interconnection networks **141** comprise multiple switch elements SE-1 **142**, SE-2 **145**, and SE-3 **148** also interconnected by multiple links. Interfaces **149** may connect via bi-directional links to line cards **139** that connect with other systems (not shown) to provide data items (e.g., packets) to be routed by packet switch **140**. Flow control information may be generated, consumed, or processed at one or more of the line cards **139**, input and output interfaces **149**, switch elements SE-1 **142**, SE-2 **145**, and SE-3 **148**, and/or other locations within packet switch **140** or the packet switching system.

FIG. 1C illustrates another exemplary operating environment and embodiment of a packet switch **160**. Packet switch **160** has a folded network topology. Packet switch **160** comprises multiple folded input and output interfaces **169** interconnected over multiple links to interconnection networks **161**, which are interconnected over multiple links returning to interfaces **169**. In one embodiment, interconnection networks **161** comprise multiple switch elements SE-1 & SE-3 **162** and SE-2 **164** also interconnected by multiple links. Interfaces **169** may connect via bi-directional links to line cards **159** which connect via ports **158** to other systems (not shown) to provide data items to be routed by packet switch **160**. Flow control information may be generated, consumed, or processed at one or more of the line cards **159**, input and output interfaces **169**, switch elements SE-1 & SE-3 **162** and SE-2 **164**, and/or other locations within packet switch **160** or the packet switching system.

FIGS. 2A–C illustrate exemplary embodiments of switching elements and/or their components in accordance with certain embodiments of the invention. FIG. 2A is a block diagram of a first stage switching element, SE-1 **200**. FIG. **213** is a block diagram of a second stage switching element SE-2 **230**. FIG. 2C is a block diagram of a third stage switching element SE-3 **260**. The invention is not limited to these or any other embodiment described herein. Rather, the invention as described herein is extensible to an unlimited

FIG. 2B

SE-3 260 generates, consumes, processes and reacts to traffic and/or flow control information as further described in detail hereinafter. Briefly first, each SE-3 260 receives packets 261 and exchanges control messages 262 over one or more links with one or more SE-2 elements (not shown) such as SE-2 230 (FIG. 2B) via SE-2 interfaces 265. In other embodiments, data packets and control messages are transmitted over a common link or links, and/or communication interfaces have a folded topology. Additionally, SE-3 260 sends packets 288 and exchanges control messages 289 over one or more links with one or more output interface elements (not shown) such as Input/Output interface 285 (FIG. 2C) via I/O interfaces 285. Control logic 271 receives control packets containing traffic and/or flow control information, and updates its flow control data structure stored in memory 272. SE-3 260 typically distributes traffic and/or flow control information to other packet switching components by sending control messages 262 and 289 as well as "piggybacking" or including traffic and/or flow control information in reserved fields of other control messages 262 and 289 (e.g., acknowledgment or clear-to-send control messages) or data packets 288 being sent. Outgoing packets 288 and control messages 289 are placed in output queues 280. Depending on the embodiment, there is an output queue 280 for each destination, for each class of service for each destination, for each next stage switching element, for each class of service for each next stage switching element, or one of many other possible configurations.

FIGS. 3A–B illustrate logical diagrams of the operation of an embodiment for collecting traffic information in tabulators, sending the collected traffic information to an accumulator, and distributing the collected flow control information to all (or a subset thereof) line cards (or ports thereof). Traffic is generally distributed across components, and/or planes of a packet switching system. For certain embodiments of packet switching systems, it is advantageous to collect indications, especially the volume and possibly location, of the distributed traffic in the various elements of a packet switching system. When an element of a packet switching system (or in some external component such as an operations system) has received this information, it can detect and react to perceived or actual traffic conditions.

FIG. 3A illustrates the one of many embodiments for collecting traffic information and generating flow control signals for a packet switching system having multiple line cards 301 each connected to an I/O interface 310. Note, the topology illustrated in FIG. 3A is that of a folded packet switch, and that each line card 301 and I/O interface 310 are shown both on the left and right side of FIG. 3A for simplicity of illustration. Also, switch elements SE-1 311 and SE-3 313 are illustrated separately; however in certain embodiments such as that illustrated in FIG. 1C, these are embodied in the same component. Moreover, other embodiments employ a different packet switch topology, such as, but not limited to a non-folded network, which provides some mechanism to convey flow control information from the output or egress portion of the packet switch back to the ingress portion.

In one embodiment as illustrated in FIGS. 3A–B, traffic information is collected by tabulators 317A–D within switching elements 312 of packet switch 300. As shown, switching elements 312 are part of the routing stage of packet switch 300. Tabulators 317A–D maintain a traffic data structure containing information to characterize the traffic level within the particular switching element. In one embodiment, each tabulator 317A–D maintains a traffic data structure to keep a count for each destination of packets within the particular switching element 312. Periodically or asynchronously, this collected traffic information is transmitted over links 329A–D to accumulators (ACC) 318A–D within next stage switching elements 313. For simplicity of illustration, traffic information in this example is shown as being sent to only one accumulator 318A, where in this and/or other embodiments, traffic information is sent to one or more accumulators 318A–D. Accumulators 318A–D maintain a data structure to accumulate the received traffic information. This accumulated information is periodically or asynchronously manipulated to determine actual, perceived or anticipated traffic conditions. Flow control information is then transmitted as illustrated by highlighted path 339 and continues over highlighted path 349 as illustrated in FIG. 3B. This flow control information can be distributed to all potential sending elements, internal or external to packet switch 300. As shown for illustrative purposes in FIG. 3B, the flow control information is broadcast to all I/O interfaces 310A–P and Line Cards 301A–P over the highlighted links. In this manner, traffic information is collected and accumulated from elements distributed within a packet switching system, with flow control messages being generated and sent to sending sources in response to the traffic conditions.

FIG. 4 illustrates a data structure 400 for storing traffic and/or flow control information. Data structure 400 comprises a table having an entry for each destination (e.g., P/O interface, line card, or port of a line card, etc.) and for each type of service supported by the packet switching system. Certain embodiments do not make a distinction between service types or only have a single class of service. As shown, data structure 400 has columns 402 corresponding to service types and rows 401 corresponding to each of the destinations, typically but not always internal to the switching system. An entry within data structure 400 is typically an integer corresponding to a packet count, although it could be another indication of traffic and/or flow control information.

FIGS. 5A–B illustrate various formats of a data structure used by various embodiments for collecting and distributing traffic and/or flow control information. FIGS. 5A–B illustrate the packet format, in which the data payload (e.g., the data fields) of the packets also illustrate a possible embodiment of the data structure (e.g., queue, stack, array, hash table) used to collect the flow control information. FIG. 5A shows one embodiment of a packet 500 having a header 501 and multiple data fields 502–504, where each data field contains an information (i.e., traffic or flow control) message. FIG. 5B shows one embodiment of a packet 510 having a header 511 and multiple data fields 512–514, where each data field contains an information (i.e., traffic or flow control) message. This embodiment uses an array of flow control messages where each data field includes the flow control information at a position within the packet (or data structure) corresponding to the source of the information. For example, data field 512 corresponds to destination 0, data field 513 corresponds to destination 1, etc.

The collection and accumulation of traffic information and distribution of flow control information is further illustrated in FIG. 6. Starting with the left of the diagram, traffic information is collected in a data structure 642 within the various SE-2 switching elements 641. Periodically or asynchronously, collected traffic information is transmitted in messages 643 to one or more SE-3 switching elements 645 where the traffic information is accumulated into data structure 648. Periodically, asynchronously, or in response to a change in a congestion or non-congestion condition, flow control messages 654 are sent to one or more packet sources 655. In response to receiving flow control messages, packet

SE-3 148, and/or other locations within packet switch 140 or the packet switching system.

FIG. 1C illustrates another exemplary operating environment and embodiment of a packet switch 160. Packet switch 160 has a folded network topology. Packet switch 160 comprises multiple folded input and output interfaces 169 interconnected over multiple links to interconnection networks 161, which are interconnected over multiple links returning to interfaces 169. In one embodiment, interconnection networks 161 comprise multiple switch elements SE-1 & SE-3 162 and SE-2 164 also interconnected by multiple links. Interfaces 169 may connect via bi-directional links to line cards 159 which connect via ports 158 to other systems (not shown) to provide data items to be routed by packet switch 160. Flow control information may be generated, consumed, or processed at one or more of the line cards 159, input and output interfaces 169, switch elements SE-1 & SE-3 162 and SE-2 164, and/or other locations within packet switch 160 or the packet switching system.

FIGs. 2A-C illustrate exemplary embodiments of switching elements and/or their components in accordance with certain embodiments of the invention. FIG. 2A is a block diagram of a first stage switching element, SE-1 200. FIG. 2B is a block diagram of a second stage switching element SE-2 230. FIG. 2C is a block diagram of a third stage switching element SE-3 260. The invention is not limited to these or any other embodiment described herein. Rather, the invention as described herein is extensible to an unlimited number of embodiments and implementations as would be understood by one skilled in the art.

Col. 4, line 63

FIG. 2A illustrates an embodiment of SE-1 200 comprising control logic and/or processor 211 (hereinafter "control logic"), memory 212, storage devices 210, I/O interfaces 205, output queues 220, SE-2 interfaces 225, and one or more internal communications mechanisms 219 (shown as a bus for illustrative purposes). In certain embodiments, control logic 211 comprises custom control circuitry for controlling the operation of SE-1 200. Memory 212 is one type of computer-readable medium, and

location, of the distributed traffic in the various elements of a packet switching system. When an element of a packet switching system (or in some external component such as an operations system) has received this information, it can detect and react to perceived or actual traffic conditions.

FIG. 3A illustrates the one of many embodiments for collecting traffic information and generating flow control signals for a packet switching system having multiple line cards 301, each connected to an I/O interface 310. Note, the topology illustrated in FIG. 3A is that of a folded packet switch, and that each line card 301 and I/O interface 310 are shown both on the left and right side of FIG. 3A for simplicity of illustration. Also, switch elements SE-1 311 and SE-3 313 are illustrated separately; however in certain embodiments such as that illustrated in FIG. 1C, these are embodied in the same component. Moreover, other embodiments employ a different packet switch topology, such as, but not limited to a non-folded network, which provides some mechanism to convey flow control information from the output or egress portion of the packet switch back to the ingress portion.

In one embodiment as illustrated in FIGs. 3A-B, traffic information is collected by tabulators 317A-D within switching elements 312 of packet switch 300. As shown, switching elements 312 are part of the routing stage of packet switch 300. Tabulators 317A-D maintain a traffic data structure containing information to characterize the traffic level within the particular switching element. In one embodiment, each tabulator 317A-D maintains a traffic data structure to keep a count for each destination of packets within the particular switching element 312. Periodically or asynchronously, this collected traffic information is transmitted over links 329A-D to accumulators (ACC) 318A-D within next stage switching elements 313. For simplicity of illustration, traffic information in this example is shown as being sent to only one accumulator 318A, where in this and/or other embodiments, traffic information is sent to one or more accumulators 318A-D. Accumulators 318A-D maintain a data structure to accumulate the received traffic information. This accumulated information is periodically or asynchronously

manipulated to determine actual, perceived or anticipated traffic conditions. Flow control information is then transmitted as illustrated by highlighted path 339 and continues over highlighted path 349 as illustrated in FIG. 3B. This flow control information can be distributed to all potential sending elements, internal or external to packet switch 300. As shown for illustrative purposes in FIG. 3B, the flow control information is broadcast to all I/O interfaces 310A-P and Line Cards 301A-P over the highlighted links. In this manner, traffic information is collected and accumulated from elements distributed within a packet switching system, with flow control messages being generated and sent to sending sources in response to the traffic conditions.

FIG. 4 illustrates a data structure 400 for storing traffic and/or flow control information. Data structure 400 comprises a table having an entry for each destination (e.g., I/O interface, line card, or port of a line card, etc.) and for each type of service supported by the packet switching system. Certain embodiments do not make a distinction between service types or only have a single class of service. As shown, data structure 400 has columns 402 corresponding to service types and rows 401 corresponding to each of the destinations, typically but not always internal to the switching system. An entry within data structure 400 is typically an integer corresponding to a packet count, although it could be another indication of traffic and/or flow control information.

Cd 8, line 26

FIGs. 5A-B illustrate various formats of a data structure used by various embodiments for collecting and distributing traffic and/or flow control information. FIGs. 5A-B illustrate the packet format, in which the data payload (e.g., the data fields) of the packets also illustrate a possible embodiment of the data structure (e.g., queue, stack, array, hash table) used to collect the flow control information. FIG. 5A shows one embodiment of a packet 500 having a header 501 and multiple data fields 502-504, where each data field contains an information (i.e., traffic or flow control) message. FIG. 5B shows one embodiment of a packet 510 having a header 511 and multiple data fields 512-514, where each data field contains an information (i.e., traffic or flow control)